

Übung 4

Rudolf Dutter & Matthias Templ
Computerstatistik (107.258) WS 2007
07. November 2007



1 Daten

Für diese Übung werden wiederum die EU-SILC Daten aus dem Jahre 2004 verwendet.

2 Übungsbeispiele

- Bsp 11 Testen Sie das Brutto-Personen-Einkommen (`pek_g`) auf Mittelwert 18147 (Signifikanzniveau 5 %). Da wir auf Grund der früheren Diskussionen vermuten, dass der wahre Mittelwert aller ungewichteten Werte größer ist, können Sie eine Richtung für den Test ausschließen. Formulieren und kommentieren Sie dabei die Hypothese, die Teststatistik, den kritischen Bereich und das Testresultat.
- Bsp 12 Verwenden Sie Variable `pek_g` (Brutto-Einkommen). Schließen Sie große Einkommen > 100000 Euro und Einkommen mit 0 Euro aus. Wie verändert sich das Testergebnis? Überlegen Sie, ob Sie einen ein- oder zweiseitigen Test formulieren sollten.
- Bsp 13 Bilden Sie 2 Gruppen Ihrer Daten durch die Variable `sex` (Geschlecht). Testen Sie auf Gleichheit der Mittelwerte (mit/ohne Ausreißer). Kann Gleichheit der Varianzen angenommen werden? Kann man daraus wiederum auf einen *Gender-Gap* schließen?
- Bsp 14 Kodieren Sie Variable `age` wie in Übung 1 und bilden Sie anschließend 4 Altersgruppen: (16, 30], (31, 45], (46, 60], (61, 85]. Testen Sie die Brutto-Einkommen gruppiert nach Altersgruppen paarweise und kommentieren Sie die 6 Vergleiche.
- Bsp 15 Bearbeiten Sie Beispiel 14 mit einem nichtparametrischen Test (z.B. Wilcoxon).

3 Abgabe

Bitte senden Sie die Ausarbeitung in Form eines *pdf*-Files (nicht mehr als 3 Seiten) mit den Resultaten (Outputs plus textliche Kommentare) und Listing des Programmcodes (Funktion) an Ihren jeweiligen Gruppenleiter bis 20. November 2007

r.dutter@tuwien.ac.at
oder
m.templ@tuwien.ac.at

Bitte den Namen des *pdf*-File wieder folgendermaßen:

name_UE_4.pdf

wobei 'name' für den Familiennamen steht.

Bitte beachten Sie, dass Sie zusätzlich Ihre gelösten Beispiele für die Übung ankreuzen müssen.